

FINAL

David A. Kidwell, Ph.D.
6125 Florence Lane
Alexandria, VA 22310
January 23, 2021

WEISBERG LAW
Matthew B. Weisberg, Esq.
Attorney ID No.: 85570
David A. Berlin, Esq.
Attorney ID No.: 314400
7 South Morton Ave.
Morton, PA 19070
610-690-0801
FAX: (610) 690-0880

Re: Dawn Kennedy v. City of Philadelphia d/b/a Philadelphia Police Department
Expert Report
Report delivered via EMAIL: dberlin@weisberglawoffices.com

Dear Mr. Berlin:

I have reviewed the additional documents relevant to the above referenced case provided by the City in December 22, 2020. I have also reviewed the rebuttal report of Dr. Leo J. Kadehjian dated November 6, 2020 in the above reference case. I take issue to his conclusions for a number of reasons, a few of which are set out below. Dr. Kadehjian's arguments have not changed my opinion that Officer Kennedy was falsely accused of drug use and that identification was due to her race in combination with the testing protocols used by the City.¹ In fact, the current analysis of the City data reinforces my opinion.

Racial Bias in Hair Testing

In my initial expert report I delved into detail how hair testing can be racially biased. I need not repeat those details here. Suffice to say that it encompasses a number of factors such as hair color, hair texture, and cosmetic treatments all closely linked to race and culture that make African-Americans and especially African-American females more susceptible to being falsely accused of drug use.

Through a subpoena, I eventually received part of the City's data set² that clearly demonstrates this bias in a statistically significant manner.

There are a number of ways to statistically analyze a given data set with varying assumptions and results. Circuit Judge's Kayatta decision in the Boston Police Case³ referenced a quip popularized by Mark Twain about statistics⁴ as "There are three kinds of lies: lies, damned lies, and statistics." I will try and parse the data in several ways to show that my conclusion that hair testing is racially bias is robust. However, it

FINAL

must be considered that if you parse the data in enough ways you can eventually convert something that is clearly statistically significant into something that is not only because the numbers become too small. I will expand upon that concept later.

The City tests their Officers in two ways. If an individual is identified to be drug tested, that individual is always tested by urinalysis and approximately 10% are given a hair test (10.0054% in the data provided⁵). First, I assume that all the hair testing is an independent event even though in a few cases the same individual tested positive several times. I still include these tests in some analyses as these tests are not repeat tests on the same sample – i.e. the sample is not just retested but a new sample is independently taken and reflects different time points of possible drug use or exposure.⁶

Individuals had self-identified their race (as provided by the City) into five categories whose proportions are broken-down in Figure 1. I was only provided the data set that reflects individuals given both a urine and hair test and my data analysis covers the time frame of January 2014 to December 2019. The raw data from which this subset was obtained consisted of ~23900 tests and is shown in Figure 1a. The set where both test were performed is in Figure 1b.^{7,8} A visual comparison of Figure 1a to Figure 1b shows that the subset analysis closely reflects the racial composition of all test performed.

Because these are random selections, these proportions should closely reflect the make-up of the whole department.⁹ There is one problem with this data set because the Federal Government considers Latino as a nationality rather than a race.^{10,11}

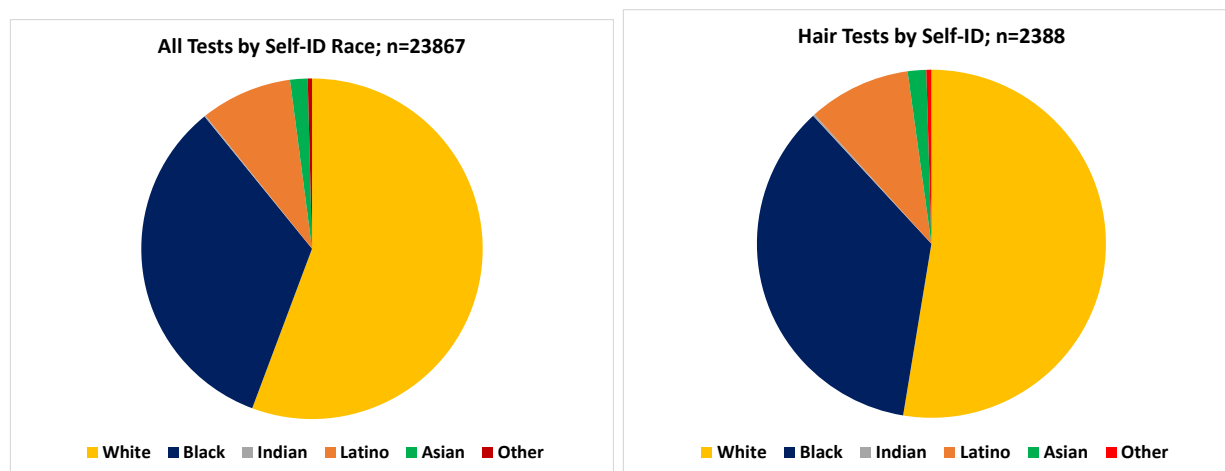


Figure 1 – The self-identified “race” of individuals in the data set analyzed. Note that this is for all tests combined and not the proportions of individuals. For example, in the Indian case, there were 32 tests reported but only 10 people with one person reflecting about 1/3 of the total tests. In the general testing population the White Officers make-up 55.7% of the tests whereas the Black Officers make-up 33.5%. In the subset of hair and urine, the White Officers make-up 52.6% whereas the Black Officers make-up 35.6%.

The proportion of individuals testing positive by hair are shown in Figure 2. Additionally, in Figure 2 is the number of individuals testing positive by urine. Both charts in comparison to Figure 1b provide a visual comparison if hair or urine testing are biased.

FINAL

Note that the urine proportions very closely mirror the general population in Figure 1a whereas the hair positive tests are skewed to the Black individuals.

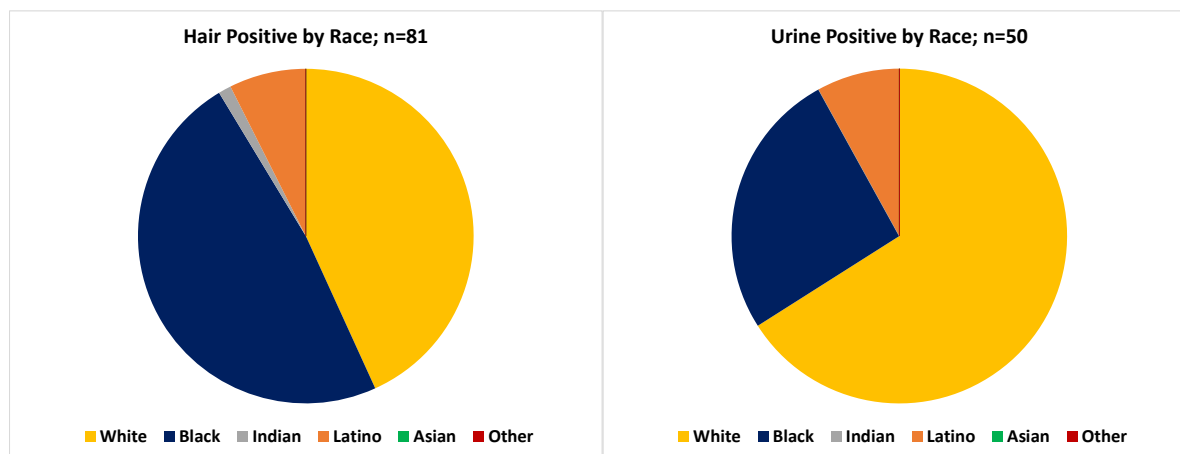


Figure 2 – Proportions of individuals who test positive for any drug by hair or urine. Because of the small number of individuals, no Asian or Other individuals are represented in either data set.

Analysis

There is no statistical difference between the two sets in Figures 1a and 1b at the 95% Confidence Interval (CI) but there is a small difference at the 99% CI level that shows a slight preference for selecting Black Officers over White Officers for hair testing.¹²

Now considering the positive drug tests. In contrast with the selection being non-biased as shown in Figure 1 (or slightly biased), the hair testing positives are statistically different at the 95% CI favoring Black Officers being positive over White Officers with an Odds ratio of 1.65 i.e. 1.65 times more likely that a Black Officer will be identified as positive than a White Officer. Like hair, urine positives have an odds ratio of 0.58 i.e. Black Officers are less likely than White Officers to be positive in urine tests but this is not statically significant due to the low number of samples (a further analysis will be provided below).

How can this be where urine is considered to be a non-biased testing process? There are several types of testing identified – random, medical evaluation, mandatory periodic test, reasonable suspicion, and reinstate. It is difficult to determine of these groups which to reasonably exclude from the analysis as with the exception of reinstate, none have a selection bias against any group. As mentioned above, we can slice the data into smaller subgroups with the proviso that even glaring differences may not become statically significant due to the small numbers.

Slicing the hair positive data into individual drugs shows that one class of drugs dominates – the opiates - at 88% of the total samples. Opiates can generally be prescribed for pain management – something Officers may be expected to incur on their

FINAL

job – or in the case of codeine for a cough. Slicing the opiates into two groups is shown in Figure 3. Of course over-use of opiates – especially medical opiates – has contributed to the opiate crisis in this country and needs to be carefully monitored.

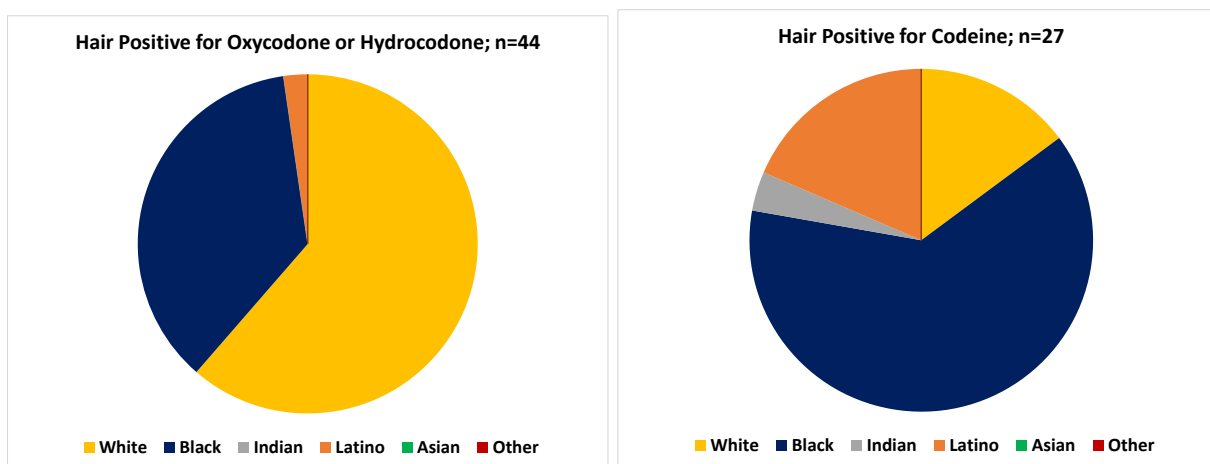


Figure 3 – Opiate Hair positives. These 71 samples represent 88% of the total 81 hair positive samples. None of these samples were reported as positive for heroin.

The identification of oxycodone or hydrocodone closely mirrors the general population whereas the codeine positives is greatly skewed to the Black Officers.

If we further slice the data in Figure 3 to those officers who were not cleared by medical review (i.e. hypothetically do not have a valid prescription for the positive opiate result) we reduce the number much further to the graphs shown in Figure 4. In the case of codeine we have reduced the number to two individuals and three hair tests. This number is too small to perform a valid statistical analysis but both individuals were Black Officers. In contrast to codeine, the other opiates show almost no bias for the Black Officers with an odds ratio 1.04 – essentially neutral.

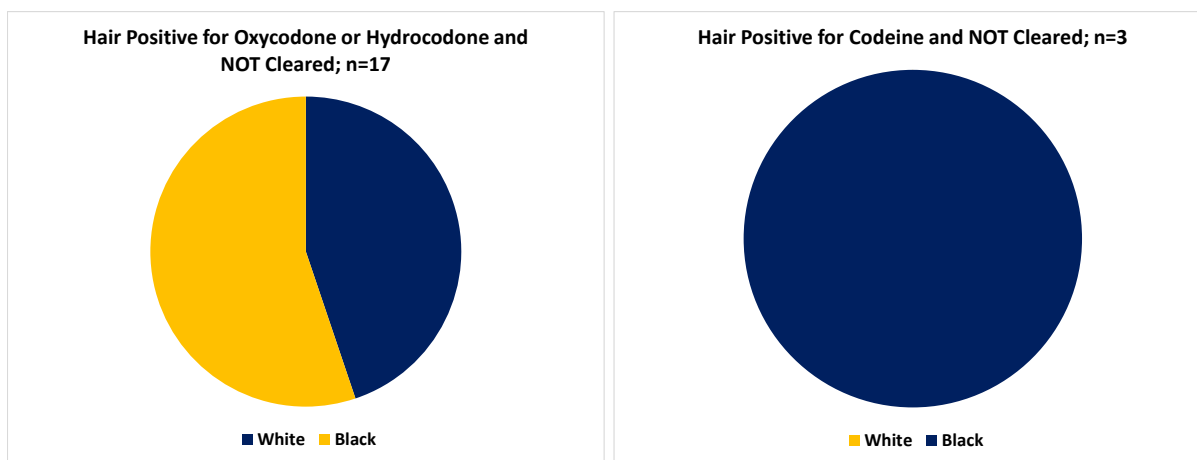


Figure 4 – Opiate hair positive without medical clearance.

FINAL

What does this sliced data mean? The Black Officers are far more likely to be identified as a codeine user but cleared. This confirms the technical literature that shows that codeine has a preference for incorporation into black hair. Thus, lower levels of use – consistent with occasional use and probably medical in nature – can be detected. That two individuals were not cleared of medical use could be a borrowed or expired prescription, likely against policy, but not that uncommon nor indicative of a true drug abuser. The fate of these two black individuals was not provided.

One more slicing of the data is to look at all drugs where the results were not cleared by medical review. This also included retests and is shown in Figure 5. This has an odds ratio of 1.82 but because of the lower numbers, this is not statically significant. However, if you include the Latino cohort as White, this increases the odds ratio to 2.14, which then is statistically significant at the 95% CI. Alternatively, if you include the Latino cohort split into equal groups into the Black and White cohorts (as suggested by U.S. Census Bureau Guidance, reference 10) the odds ratio drops back to 1.75 and the results are now no longer statistically significant. This is a perfect example of slicing data too finely and may be a good illustration of Mark Twain's quip.

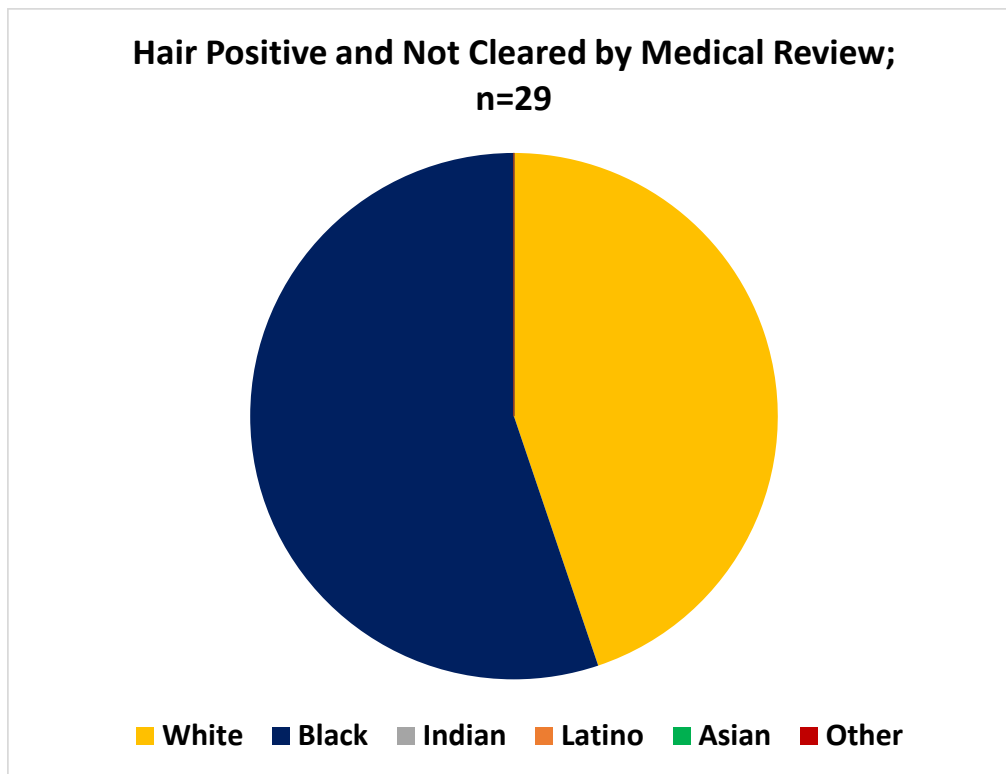


Figure 5 – Hair positive samples, not cleared by medical review. One of these individuals is from the present case.

As to the instant case, only five Officers have been identified as marijuana positive by any testing.¹³ Of those five, one was identified by urine alone, one by hair and urine, and three by hair alone. The hair alone cases showed two Black Officers (including the

FINAL

present case) and one White Officer. Of course even a 2:1 odds ratio is insufficient with only three people for statistical analysis.

One last slicing of the data, where I consider only those individuals who were randomly selected rather than for some reason.¹⁴ In this case the number of hair positives drops to 14 but the odds ratio increases to 14.8, which is statistically significant even with these low numbers at the 95% CI. A graph of the results is depicted in Figure 6. Officer Kennedy is part of this group and maybe that is the best comparison of all.

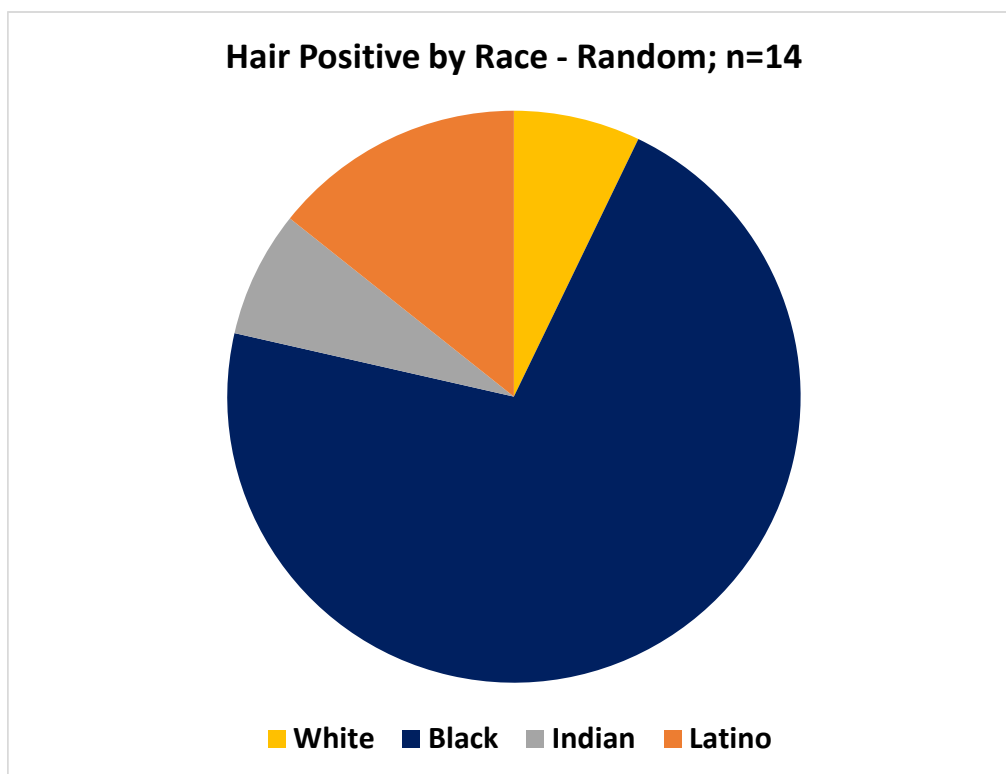


Figure 6 – Individuals who were hair positive for any drug and who were selected at random.

Urine Positives

As mentioned above, urine positives have an odds ratio of 0.58 i.e. Black Officers are less likely than White Officers to be positive in urine tests. Although this is not statistically significant due to the low number of samples, we can slice the data even further. If you just consider the urine positives with no medical clearance, the number drops to 18 but the odds ratio increases to 0.89. However, again due to the low numbers, this is not statistically significant i.e. urine does not show a bias. The data is skewed towards non-black officers because seven of these positives were for steroid use where the sample was taken because of reasonable suspicion i.e. these were not random samples.

FINAL

Rebuttal to report of Leo J. Kadehjian dated November 6, 2020

As summarized in the introduction, Dr. Kadehjian analysis provides no further information nor does his rebuttal change my conclusions. Many of his points are very repetitious as if repetition of a claim or opinion makes it true. However a few points need to be made at this time.

Dr. Kadehjian appears to focus on my examples with cocaine where I show that hair testing can confuse use with exposure in this well-studied drug. Dr. Kadehjian says nothing about my THC references or analogy of cocaine to THC. He should review the studies of drugs in children, some of which I cite, for examples of external exposure.

As to cut-offs, Dr. Kadehjian needs to review the history of hair testing and some of the published data. The cut-offs and procedures have changed dramatically over the past decades. He need not look any further than the 2020 SAMHSA proposed guidelines for THC where the cut-off is half that used in the instant case. I was involved in the discussions that prompted some of these changes over the past decades. They are arbitrary and based partly on what the laboratories can accomplish and partly on what is necessary to get a sufficient number of positive samples. Very infrequently were false positives considered in setting cut-off levels. Certainly Dr. Kadehjian is aware of ROC curves in selection of optimum testing cut-offs. They are rarely used in hair testing because you have few ways of determining the false positives from the true positives. That is where our pioneering data studying children could have been useful in setting cut-offs. Unfortunately, there was no practical cut-off to distinguish users from non-users in that data set for cocaine. As a compromise, I proposed the 10% BE rule for cocaine. It was adopted for about 6 months by some commercial laboratories but then quickly abandoned because there were insufficient positives under that rule. Additionally, perhaps Dr. Kadehjian is perhaps unaware of changes that at least two commercial laboratories have made in their established procedures (some followed for decades) because they were not getting a sufficient positive rate. Some of this information is public in the filings on the Boston Case. Dr. Kadehjian appears to believe that guidelines proposed by learned societies are followed to the letter. That does not even happen by founding members of those societies. Unlike some drugs and medical procedures cleared by the FDA, there is no force of law in many forensic procedures. It is up to the Courts to wade through the scientific information and decide if the science is valid and is being fairly and uniformly applied. Only if sufficient precedence is developed will the equivalent of the force of law be in place because Courts will not allow faulty science to be introduced. Unfortunately, this takes time and money for those cases to be developed and frankly the Courts are not always the best place to develop this science.

As to lengths of hair. Officer Kennedy's hair was tested at too long of length. The laboratory had protocols to test hair cut to 1.5". That was not done in her case and thus subjected her sample to, in my opinion, different treatment from that received by other individuals similarly situated. Dr. Kadehjian need not look any further than the Omega litigation package that shows that Officer Kennedy's hair was the longest tested in that

FINAL

batch (where length was provided, page 12 of 82, for his reference). One can speculate as to why this was the case. I will not do that here but this is a form of purposeful discrimination in not following protocol. In my experience, African-American female hair is very curly and difficult and time consuming to align and cut to length if is not cut during the collection process, as this sample was not. Why the Omega collection process did not have a requirement to cut the hair to a defined length while on the head is unclear. I do note that USDTL, the laboratory where Officer Kennedy sent a second sample, did have the requirement that the collection facility cut the hair to length and it appeared to have been met in the instant case.

I am pleased that Dr. Kadehjian acknowledges that the current SAMHSA guidelines were published after my report and thus I could not discuss them as I am not clairvoyant. I did read those guidelines when they came out, and as an expert in this field, I commented extensively on selected parts of the guidelines (not all due to time limitations). Those comments are available for public review. I did not find any comments by Dr. Kadehjian on the web and so apparently he did not wish to add his expertise to help SAMHSA further hair testing.¹⁵

I understand that Federal regulations are difficult to read and understand. I also find that to be a problem. However, Dr. Kadehjian is just plain wrong in hair length required in those proposed guidelines. Instead of cutting the hair during collection, SAMHSA is requiring the laboratories to cut the hair to 1" length during analysis. This is noted as:

Section 11.10 What amount of hair is tested?

The laboratory prepares an aliquot of the hair specimen of the specified weight needed for the test. If the root end is identified, the laboratory uses the first one inch of the hair from the root end. (Underlined for emphasis)

To understand this section one must also delve into the collection process where hair longer than 1" has the root identified and thus 1" is the maximum length.¹⁶ If this procedure goes into practice as written, it will cause problems with African-American female hair because of its curly nature.

I like the way Dr. Kadehjian repeatedly and repeatedly and repeatedly puts the word facts in quotes. I note that Dr. Kadehjian appears to confuse 'facts' i.e. observations with speculation i.e. hypothesis in much of his analysis of the technical literature.

Dr. Kadehjian is glaring inconsistent in his statement:

.... Dr. Kidwell also addresses the financial costs of testing vs. the costs of replacing an officer. But he ignores the costs of allowing an officer who has violated the Department drug use policies to continue, potentially putting the Department and the public at risk.

I have not ignored the cost of allowing an office to continue to violate the Department drug use policies. In fact, using hair testing as a detection method does just that! Dr.

FINAL

Kadehjian has stated that drugs do not appear in hair immediately (which I do not agree with) but let's go with it. So an inebriated officer would be negative by hair testing because the 10+ days have not passed before the hair with the drug grows from the scalp. In fact, as hair testing records back 90+ days, was the officer problematic historically? What does possible historic drug use have to do with current job performance? By identifying individuals who are potential drug users, one can apply technology that really determines if that individual is a drug user. Urinalysis is such a technology. I outlined in my initial report how to use urinalysis in conjunction with hair testing in an enlightened manner that meets or actually improves the current system, and does it in a way that is not racially biased. As an aside, SAMHSA appears to suggest using my approach or something akin to it. Additionally as currently practiced, hair testing does not detect steroid use whereas seven officers in this subset were steroid positive (two were positive for other drugs via hair testing)! Steroids have been linked to aggression.

Conclusions

Research suggests that African Americans are disparately impacted by false accusations of drug use when hair testing for drugs is employed. Genetics and cultural differences in how hair is cosmetically treated provides African Americans with hair that is more susceptible to contamination from the environment and thereby generating false accusations of drug use from mere drug exposure.

Statistical analysis showed that hair testing by the City had a disparate impact on African Americans. Officer Kennedy is a select member of that group who was randomly chosen for hair testing and was positive. This random group had an odds ratio of 14.8 (i.e. 14.8 times more black individuals (in proportion to their population) were selected than white individuals), which was statistically significant at the 95% CI.

Also, testing of too long a hair sample subjected Officer Kennedy to different treatment from that received by other individuals similarly situated.

I have rendered all the above opinions and conclusions to a reasonable degree of scientific certainty; and, in forming my opinion and conclusions, I have relied on documents and scientific techniques that are reasonable, customary, and necessary in my field of drug-testing expertise.

FINAL

References and Notes

¹ This report is being given in my capacity as a private citizen. I am being compensated for writing this report. All parts were written by myself. The opinions contained herein represent my own views and do not necessarily represent the opinions or policies of the Department of the Navy or of the U.S. Government.

² Numbered City 1031 – City 1075.

³ Hon. G.A. O'Toole, Jr., U.S. District Judge, United States Court of Appeals for the First Circuit, No. 12-2280, Ronnie Jones, et al. vs. City of Boston, Boston Police Department; Edward Davis, Commissioner of the Boston Police Department, Appeal From the United States District Court for the District of Massachusetts before Judges Torruella, Howard, and Kayatta, 05/07/2014.

⁴ Judge Kayatta, in discussion of various ways to parse data, just referred to the quip without actually stating it verbatim.

⁵ Yes, I remember my lessons on significant figures and round this to 10%.

⁶ Likewise, an individual who tested positive, then negative would still be considered positive as would an individual who tested negative, then positive. All these tests are considered independent events.

⁷ However, as indicated in the initial discussion, the number of urine positives is limited to 46 Black and White individuals. More data parsing could be done if the whole data set were to be provided.

⁸ The City did not provide the race of the individuals in the December data set. Instead, I needed to correlate the date of test and name with a master sheet to determine the self-identified race. Six individuals were identified in the hair subset who could not be found in the master sheet. These six individuals were all negative and not included in the data analysis.

⁹ Even if these proportions did not reflect the Department as a whole, they do represent the individuals similarly situated – i.e. being drug tested.

¹⁰ The "U.S. Census Bureau Guidance on the Presentation and Comparison of Race and Hispanic Origin Data" found at: [Comparing Race and Hispanic Origin \(census.gov\)](https://www.census.gov/data/tables/2010/special/hespan/comp-race-hispanic.html) is not entirely helpful on reclassifying the Latino group self-identified in the City data into the Black and White categories.

¹¹ Also, I am concerned that the term "Indian" in the data set really means individuals from India rather than Native Americans. Individuals from India are generally classified as Asian. However, these numbers are low and will be ignored in the data analysis.

¹² Magdalena Szumilas, "Explaining Odds Ratios", *J Can Acad Child Adolesc Psychiatry*, 19:3, (2010) 227.

¹³ One White Officer was identified as positive by urine for steroids but positive by hair for marijuana. However, the data for urine or hair is restricted to listing one drug. I have assumed that the urine would have also been positive for marijuana. Hair is not tested for steroids.

¹⁴ When performing statistics one must assume a distribution for the data. This is generally assumed to be a Gaussian distribution unless shown otherwise. Gaussian distributions are most frequently obtained by a random selection process.

¹⁵ The public comments are available at: [Regulations.gov - Docket Folder Summary](https://www.regulations.gov/docket?cid=1568488). I searched this website for Dr. Kadehjian's name and it was not found. Some comments are anonymous. Perhaps he wished to comment in that manner. I also reviewed some of the 213 comments where the title appeared interesting.

¹⁶ 1.5" comes from an approximation of the growth rate, which is normally given as a range of 0.8-1.3 cm/month. As is normal with humans, there are outliers even in this broad of range. Thus, 3 cm or 1.18" corresponds to 3 months of growth. SAMHSA appears to be using the lower growth rate to cover the 3 month period as $0.8 \times 3 \sim 1$ ". This is conservative and reasonable to treat individuals fairly.